

KHOA HỌC XÃ HỘI

Đánh giá học tập bậc đại học trong thời AI: Từ chống gian lận sang thiết kế minh chứng tư duy

Vũ Thị Thu Thanh*,¹

¹Viện Khoa học xã hội vùng Nam Bộ, Viện Hàn lâm Khoa học xã hội Việt Nam

*Corresponding author:

Vũ Thị Thu Thanh - Viện Khoa học xã hội vùng Nam Bộ, Viện Hàn lâm Khoa học xã hội Việt Nam.

Email: thuthanhkxh@gmail.com

Received: 13/03/2026. **Revised:** 01/05/2026. **Accepted:** 08/05/2026. **DOI:** [10.53901/tjs.2026.v01.issue02.art11](https://doi.org/10.53901/tjs.2026.v01.issue02.art11)

Tóm tắt

Đánh giá giáo dục đại học, đặc biệt ở các môn khoa học xã hội và nhân văn, đang đối mặt với một sự khủng hoảng, thách thức khi các công cụ AI tạo sinh (GenAI) đang ngày càng phát triển và phổ biến. Bởi lẽ, việc đánh giá cũ dựa trên một giả định rằng sản phẩm văn bản nộp cuối kỳ là bằng chứng đáng tin cậy cho năng lực tư duy và hiểu biết của người học. Tuy nhiên, giả định nền tảng này bị suy yếu khi GenAI có thể tạo ra sản phẩm hoàn chỉnh trong một thời gian ngắn. Từ đó, việc đánh giá rơi vào vòng luẩn quẩn: cảm đoán, đối phó và truy vết. Bài viết lập luận rằng thay vì tập trung ngăn chặn AI, cần tái thiết kế đánh giá để đo đúng năng lực cốt lõi: tư duy học thuật, năng lực lựa chọn và trách nhiệm giải trình. Để đảm bảo tính hợp lệ của việc đánh giá, bài viết đề xuất chuyển hệ mục tiêu từ nhấn mạnh vào sản phẩm văn bản đầu ra sang minh chứng tư duy. Trên cơ sở đó, bài viết đề xuất bốn mô hình đánh giá phù hợp thời AI và quy tắc liên chính học thuật dựa trên minh bạch sử dụng AI cùng quy trách nhiệm cho người học.

Từ khoá: Trí tuệ nhân tạo tạo sinh; đánh giá học tập; minh chứng tư duy; liên chính học thuật; giáo dục đại học.

Đặt vấn đề

Trí tuệ nhân tạo tạo sinh (Generative Artificial Intelligence – GenAI) là một nhánh quan trọng của trí tuệ nhân tạo, tập trung vào việc phát triển các mô hình có khả năng học các mẫu hình thống kê từ dữ liệu lớn để tạo ra nội dung mới dưới dạng văn bản, hình ảnh, âm thanh hoặc các dạng dữ liệu khác. Sự phát triển và phổ biến nhanh chóng của các công nghệ này đang trở thành chủ đề trung tâm trong các cuộc thảo luận về đào tạo và đánh giá ở giáo dục đại học [1–5]. Mặc dù GenAI mang lại những cách tiếp cận mới mẻ trong giảng dạy, đánh giá và nâng cao mức độ hứng thú cũng như sự tham gia của người học, nhưng việc lạm dụng quá mức công nghệ này đặt ra những thách thức đáng kể cho cả giảng viên và sinh viên [6, 7].

Trong giáo dục, việc đánh giá kết quả học tập là một khâu then chốt, không chỉ đơn thuần là công cụ đo lường mà còn là nhân tố quyết định chất lượng và thành quả học tập của người học [1, 8]. Khi được thiết kế tối ưu, hoạt động đánh giá đóng vai trò như một công cụ định hướng giúp sinh viên thấu hiểu năng lực bản thân, từ đó cải thiện đáng kể hiệu quả và trải nghiệm học tập [9]. Trước thời kỳ GenAI, giáo dục đã dựa trên một giả định ngầm định rằng nếu một học

viên có thể nộp một sản phẩm văn bản đạt tiêu chí đề ra thì đó được xem là bằng chứng cho thấy người học đã tư duy, đã hiểu biết và đã đạt được kết quả học tập. Giả định này hợp lý vì sản phẩm văn bản đầu ra thường là kết quả của một quá trình dài bao gồm đọc tài liệu, suy nghĩ phê phán, viết nháp, chỉnh sửa và tái cấu trúc. Ngày nay, với sự hỗ trợ của các mô hình GenAI, một sản phẩm văn bản hoàn chỉnh, có chất lượng cao có thể được tạo ra chỉ trong thời gian ngắn. Điều này làm suy giảm giá trị của việc đánh giá năng lực thực sự dựa trên công việc tự thực hiện với công việc được hỗ trợ một phần và hỗ trợ hoàn toàn bởi AI. Do đó, giáo dục đại học không chỉ đơn thuần là việc tích hợp GenAI vào lớp học, mà còn đòi hỏi một sự xem xét lại một cách nghiêm túc về giả định nền tảng của đánh giá học tập, cũng như việc xác định những gì có thể được coi là bằng chứng đáng tin cậy và xác thực cho quá trình học tập của người học.

Trước sự phổ biến và lạm dụng GenAI của người học, phản ứng chung của nhiều cơ sở giáo dục hiện nay là cấm sử dụng GenAI kết hợp với việc tăng cường các công cụ phát hiện gian lận. Tuy nhiên, cách tiếp cận này tồn tại những bất cập lớn. Thứ nhất, nó khiến nhà trường rơi vào thế đuổi bắt công nghệ: các công cụ phát hiện luôn chậm chân hơn so với sự tiến bộ của GenAI, trong khi hành vi lạm dụng ngày càng đa

dạng và tinh vi. Thứ hai, biện pháp cảm đoán không chạm đến gốc rễ vấn đề vì gần như AI đã đi sâu vào các hoạt động xã hội nên việc ngăn chặn tuyệt đối là bất khả thi mà cần nhìn nhận đó là khủng hoảng của mô hình đánh giá dựa chủ yếu vào sản phẩm cuối cùng vốn đã không phản ánh năng lực tư duy thực sự của người học trong thời đại AI. Vì vậy, vấn đề cốt lõi không nằm ở việc làm sao ngăn chặn GenAI khỏi lớp học, mà là làm sao thiết kế hệ thống đánh giá để đo lường chính xác những năng lực cốt lõi mà trí tuệ nhân tạo khó có thể thay thế hoàn toàn. Nói cách khác, thay vì coi GenAI như một yếu tố phá hoại, gây gian lận, cần xem nó như một chất xúc tác thúc đẩy hệ thống đánh giá quay về bản chất giáo dục đúng đắn: tập trung đo lường năng lực tư duy học thuật, khả năng lập luận, năng lực lựa chọn/phán đoán và trách nhiệm học thuật. Như Xia và các cộng sự chỉ ra, GenAI đang thay đổi đánh giá ở ba cấp độ (người học, giảng viên và cơ sở giáo dục), đòi hỏi tái thiết kế để thúc đẩy học tập tự điều chỉnh và các kỹ năng cao cấp hơn, phù hợp với hướng tiếp cận dựa trên minh chứng tư duy [3].

Từ góc nhìn này, việc đánh giá giáo dục cần một sự chuyển trục mục tiêu: từ việc đánh giá chủ yếu dựa vào đầu ra là sản phẩm văn bản sang việc đánh giá dựa trên minh chứng tư duy. Minh chứng ở đây không chỉ giới hạn ở sản phẩm nộp cuối cùng, mà bao gồm toàn bộ dấu vết học tập có thể kiểm chứng và truy vết, chẳng hạn như cách lựa chọn câu hỏi nghiên cứu, cấu trúc lập luận, các phiên bản nháp, quyết định sửa đổi, cách đối chiếu và kiểm chứng nguồn tài liệu, cách phản biện lại chính mình, cũng như cách người học giải thích lý do chọn một hướng tiếp cận thay vì hướng khác. Với cách tiếp cận này, GenAI sẽ không còn là lối đi tắt để qua môn; ngược lại, nó trở thành một công cụ buộc người học phải thể hiện rõ ràng năng lực kiểm soát, làm chủ và phê phán công cụ đó.

Mặc dù chủ đề tích hợp GenAI vào đánh giá đại học đã được nhiều nghiên cứu quốc tế đề cập, các tiếp cận hiện hành vẫn để lại ba khoảng trống đáng kể. Thứ nhất, các khung tích hợp như AIAS hay khung của Khlaif và các cộng sự tập trung phần lớn vào quy mô và mức độ cho phép sử dụng AI mà chưa đi sâu vào hình thái và quy trình minh chứng cụ thể để giảng viên triển khai trong từng kiểu học phân [4, 10]. Thứ hai, các khuyến nghị về đánh giá hướng vào quá trình của O'Sullivan và các cộng sự hay Sallai và các cộng sự mới ở cấp nguyên tắc, chưa được hệ thống hóa thành các mô hình đánh giá có quy trình triển khai và thang đánh giá chi tiết [5, 11]. Thứ ba, hầu hết tài liệu hiện hành xuất phát từ ngữ cảnh giáo dục Âu - Mỹ, chưa bàn đến các điều kiện thực tiễn tại Việt Nam - nơi quy mô lớp lớn, văn hóa đánh giá nặng về điểm số tổng kết và năng lực số của giảng viên còn không đồng đều.

Trên cơ sở các khoảng trống nêu trên, bài viết đặt ra ba đóng góp chính. Về phương diện lý luận, bài viết hệ thống hóa nguyên lý chuyển hệ mục tiêu từ sản phẩm văn bản sang minh chứng tư duy, cụ thể hóa

thành hai thuộc tính cốt lõi (tính truy vết và tính giải thích được) cùng ba năng lực then chốt (lập luận, lựa chọn - phán đoán, trách nhiệm học thuật). Về phương diện mô hình ứng dụng, bài viết đề xuất bốn mô hình đánh giá phân hóa theo bối cảnh sư phạm, mỗi mô hình có quy trình triển khai, thang đánh giá đa tiêu chí và ánh xạ với các cấp độ tương ứng của khung AIAS. Về phương diện thực tiễn, bài viết bước đầu thảo luận điều kiện triển khai trong giáo dục đại học Việt Nam, tạo cơ sở cho các nghiên cứu thí điểm tiếp theo.

Phương pháp nghiên cứu

Bài viết sử dụng phương pháp phân tích, tổng hợp tài liệu định tính có cấu trúc (structured qualitative literature analysis) nhằm hệ thống hóa khung lý luận và đề xuất mô hình đánh giá phù hợp với bối cảnh GenAI. Quy trình thực hiện gồm ba bước:

Bước 1: Lựa chọn tài liệu. Tài liệu được tuyển chọn theo bốn tiêu chí: (i) tính cập nhật, ưu tiên các nghiên cứu xuất bản từ năm 2019 đến 2026, đặc biệt sau cột mốc xuất hiện ChatGPT (cuối năm 2022); (ii) tính chuyên sâu, ưu tiên các bài đăng trên tạp chí khoa học bình duyệt và báo cáo của các tổ chức giáo dục quốc tế uy tín; (iii) tính liên quan trực tiếp đến hai trục chủ đề là đánh giá học tập đại học và tích hợp GenAI vào đánh giá; (iv) tính bổ trợ về quan điểm, gồm cả nguồn ủng hộ tích hợp AI [3, 5, 10-14] và các nguồn cảnh báo rủi ro [6, 7, 15] để đảm bảo cân bằng góc nhìn.

Bước 2: Phân tích và đối sánh. Các tài liệu được trích xuất theo bốn nhóm thông tin: (i) giả định nền tảng của đánh giá đại học; (ii) tác động của GenAI tới giả định đó; (iii) các nguyên tắc thiết kế đánh giá thời AI; (iv) các hình thái minh chứng học tập được khuyến nghị. Kỹ thuật đối sánh chéo (cross-comparison) được áp dụng để xác định những điểm hội tụ, ví dụ sự nhất trí giữa O'Sullivan và các cộng sự [5], Sallai và các cộng sự [11] và Page và các cộng sự [12] về tính ưu việt của đánh giá hướng vào quá trình, cũng như các điểm phân kỳ giữa AIAS [10, 13, 14] và các khung khác về mức độ cho phép sử dụng AI.

Bước 3: Khái quát hóa thành mô hình. Trên cơ sở khung lý luận đã hệ thống hóa, bốn mô hình đánh giá được hình thành theo logic giao thoa hai trục: trục thời gian của minh chứng (quá trình ↔ kết quả tổng hợp) và trục hình thái minh chứng (văn bản tích lũy-đối thoại- sản phẩm tích hợp). Mỗi mô hình được mô tả theo cấu trúc thống nhất gồm: bối cảnh sư phạm phù hợp, quy trình triển khai, thang đánh giá đa tiêu chí, và ánh xạ với cấp độ tương ứng trong khung AIAS [10, 13, 14]. Việc thiết kế bốn mô hình theo cấu trúc đồng nhất nhằm đảm bảo tính khả triển khai (deployability) cho giảng viên và tính so sánh được giữa các mô hình.

Chuyển đổi mục tiêu: Từ đầu ra là sản phẩm văn bản sang minh chứng tư duy

Khi sự phát triển và phổ biến của các mô hình ngôn ngữ lớn có thể tạo ra văn bản đúng cấu trúc, mạch lạc, thậm chí mang vẻ sâu sắc chỉ trong thời gian rất ngắn, hệ quả là đầu ra văn bản không còn phản ánh chính xác năng lực thực của người học [15]. Do đó, vấn đề không chỉ nằm ở gian lận học thuật theo nghĩa đạo đức, mà cốt lõi là sự suy giảm tính hợp lệ của công cụ đánh giá khi sản phẩm văn bản nộp không còn là bằng chứng đủ tin cậy cho những gì nhà trường muốn đo lường. Nếu hệ thống tiếp tục lấy đầu ra là văn bản làm trung tâm, nó dễ rơi vào tình trạng vừa căng thẳng vừa bất lực: càng kiểm soát càng kích hoạt đối phó; càng siết chặt càng làm nghèo trải nghiệm học tập; càng lệ thuộc vào việc kiểm lỗi càng khiến mục tiêu giáo dục lệch khỏi việc phát triển năng lực con người trong thời đại mới.

Để thoát khỏi vòng luẩn quẩn: cảm đoán, đối phó, truy vết, cần một chuyển đổi mục tiêu mang tính phương pháp luận trong đánh giá. Thay vì đặt câu hỏi truyền thống “văn bản này có phải do người học tự viết không?”, hệ thống đánh giá cần chuyển sang trọng tâm cốt lõi hơn: đánh giá quá trình học tập chính nó thông qua các bản nháp, phân phản ánh và những cải tiến lặp lại, thay vì chỉ tập trung vào sản phẩm cuối cùng. Hay nói cách khác, cần chuyển sang câu hỏi then chốt: “văn bản này tiết lộ quá trình tư duy của người học như thế nào, và đâu là những minh chứng cụ thể cho quá trình ấy?”, với yêu cầu sinh viên phải giải trình rõ lý do chấp nhận, chỉnh sửa hay bác bỏ gợi ý từ AI, từ đó thể hiện sự tự chủ trí tuệ phân biệt giữa học tập thực sự của con người và việc sao chép cơ học [5].

Như Xia và các cộng sự [3] tổng quan về đánh giá kết quả học tập bậc đại học cho thấy việc tích hợp GenAI vào đánh giá học tập (xuất hiện trong 11/34 nghiên cứu) là một trong ba xu hướng chính. Điều đó có nghĩa là chấp nhận việc sử dụng các công cụ GenAI của sinh viên nhưng thay vì tiếp nhận thụ động, cần hướng người học phải làm việc cùng AI thông qua việc phân biện các giải pháp từ ChatGPT hay sử dụng công cụ tạo hình ảnh để minh chứng cho khả năng sáng tạo [16, 17]. Những chuyển dịch này không chỉ đổi mới hình thức kiểm tra mà còn là tiền đề để các nhà nghiên cứu khai thác sâu hơn tiềm năng của GenAI trong giáo dục. Sự chuyển hướng này định hình một hệ mục tiêu mới theo thứ tự logic, từ đầu ra văn bản sang quá trình tư duy và cuối cùng là minh chứng học tập. Trong sự dịch chuyển đánh giá mới, văn bản không bị loại bỏ nhưng nó không còn là đích đến duy nhất mà trở thành một phần trong chuỗi minh chứng rộng lớn hơn. Các thao tác cốt lõi nhằm xác định và đo lường người học bao gồm: người học hiểu và xác định trọng tâm của yêu cầu đánh giá như thế nào; biết lựa chọn luận điểm, loại bỏ ý tưởng không phù hợp, kèm theo lý do rõ ràng; biết thiết lập mạch suy luận logic, sử dụng bằng chứng, phản biện, và giải thích mối quan hệ nhân quả hoặc khái

niệm; biết đối chiếu và kiểm chứng bằng cách so sánh các nguồn tài liệu, nhận diện mâu thuẫn, phát hiện sai sót, và thực hiện sửa chữa; cuối cùng là khả năng tự phản tư, một kỹ năng quan trọng để tránh thiên kiến xác nhận khi mà các GenAI rất dễ nắm bắt và trả lời theo đúng ý của người dùng. Kỹ năng này bao gồm việc nhận diện giới hạn và thiên kiến của bản thân cũng như điểm yếu trong lập luận và đề tự điều hướng cách cải thiện và học hỏi.

Khi hệ thống đánh giá tập trung vào các thao tác tư duy cốt lõi này, GenAI khó có thể trở thành người làm thay để người học vượt qua ngưỡng điểm số một cách thụ động. Thay vào đó, sinh viên buộc phải chứng minh khả năng kiểm soát và làm chủ toàn bộ quá trình suy nghĩ của mình, chứ không chỉ nộp một sản phẩm hoàn chỉnh. Như tài liệu hướng dẫn của O’Sullivan và các cộng sự [5] nhấn mạnh trọng tâm của hệ thống đánh giá mới chính là minh chứng học tập: một tập hợp các dấu vết cụ thể cùng lời giải trình thể hiện mức độ đạt được mục tiêu học tập. Khác hẳn với bài làm cuối kỳ truyền thống (chỉ tập trung vào sản phẩm cuối), minh chứng tư duy nhấn mạnh hai thuộc tính then chốt:

- Thứ nhất là tính truy vết, cho phép người đánh giá lần theo tiến trình hình thành ý tưởng và sản phẩm một cách tự nhiên. Các tổ chức nên ưu tiên trọng số chính cho bằng chứng quá trình, chẳng hạn như lịch sử bản nháp, nhật ký phiên bản và đường dẫn nghiên cứu.
- Thứ hai là tính giải thích được, người học phải có khả năng giải trình lý do đằng sau các lựa chọn của mình, thông qua việc trình bày công việc của họ bằng các tài liệu như ghi chú lập kế hoạch, bản nháp, lịch sử câu lệnh, giải thích miệng và bình luận phản ánh.

Những yếu tố này đảm bảo rằng GenAI chỉ đóng vai trò hỗ trợ, còn trách nhiệm cuối cùng về lập luận, quyết định và cam kết học thuật vẫn thuộc về người học, giúp giảm thiểu rủi ro lạm dụng công cụ một cách thụ động. Các dạng minh chứng tư duy bao gồm:

- Bản nháp theo phiên bản (phiên bản 1, 2, 3) kèm ghi chú về các thay đổi và lý do sửa chữa;
- Nhật ký ghi chép cho thấy các quyết định của người học như ghi chép các lựa chọn quan trọng (chọn A vì..., bỏ B vì...);
- Bảng so sánh và đối chiếu nguồn bao gồm phân tích điểm tương đồng/khác biệt, đánh giá mức độ tin cậy, và lý do ưu tiên nguồn;
- Bản đồ lập luận bao gồm trình bày luận điểm - bằng chứng - phản biện - kết luận;
- Bản tự đánh giá bao gồm phản ánh những gì đã học được, sai sót ở đâu, và cách khắc phục;
- Bản khai báo sử dụng AI bao gồm nêu rõ bước sử dụng, mục đích, và mức độ chỉnh sửa sau khi áp dụng AI.

Mình chứng tư duy không nhằm mục đích gây khó người học hay tăng gánh nặng cơ học cho giảng viên. Ngược lại, nó giúp giảng viên đánh giá chính xác những năng lực cần đo lường, đồng thời giảm thiểu tình trạng phải dự đoán về năng lực thực chất của người học. Khi chuyển trọng tâm sang mình chứng tư duy, các tiêu chí đánh giá cũng được điều chỉnh tương ứng. Thay vì ưu tiên sự trơn tru của văn phong hay độ hoàn chỉnh bề mặt của sản phẩm, hệ thống mới tập trung vào ba năng lực cốt lõi khó bị GenAI thay thế hoàn toàn. Đầu tiên là năng lực lập luận, bao gồm xây dựng luận điểm chặt chẽ, sử dụng bằng chứng phù hợp, phản biện hiệu quả, và kết nối các ý tưởng một cách logic; thứ hai là năng lực lựa chọn và phán đoán, là khả năng chọn phương pháp tiếp cận, nguồn tài liệu, cấu trúc bài viết, và mức độ sử dụng AI một cách có lý do. Sallai và các cộng sự [11] đề xuất cần tách biệt rõ ràng quá trình học tập khỏi sản phẩm được đánh giá, đồng thời thiết kế các bài tập có yếu tố liên tục trước khi nộp, yêu cầu người học ghi chép lại sự phát triển của sản phẩm cuối; và thứ ba là năng lực chịu trách nhiệm học thuật, bao gồm minh bạch về hỗ trợ công nghệ, trích dẫn đúng chuẩn, kiểm chứng thông tin, và sẵn sàng bảo vệ các lựa chọn của mình. Page và các cộng sự [12] cho rằng AI chỉ giỏi cung cấp những câu trả lời trong thang nhận thức thấp của Bloom, thay vì cảm đoán giáo dục cần tái định nghĩa phương thức đánh giá thông qua việc lồng ghép công cụ này một cách có ý nghĩa nhằm thúc đẩy sự tương tác, phản hồi và tính thực tế trong quy trình học tập, nhấn mạnh rằng kỹ năng phân tích phê phán nội dung do AI tạo ra cùng với việc khám phá cách sử dụng công nghệ một cách đạo đức chính là những kỹ năng then chốt mà sinh viên cần để thành công trong bối cảnh hiện nay. Mặc dù các công cụ GenAI có thể làm cho một số nhiệm vụ trở nên nhanh chóng hơn, chúng vẫn không thể thay thế kiến thức chuyên sâu về môn học và khả năng phán đoán của chuyên gia con người [11].

Bốn mô hình đánh giá phù hợp trong giáo dục thời đại AI

Khung Đánh giá Trí tuệ Nhân tạo (AI Assessment Scale - AIAS)

Bài viết của Perkins và các cộng sự [10], được xuất bản trên Journal of University Teaching and Learning Practice, giới thiệu Khung Đánh giá trí tuệ nhân tạo (AI Assessment Scale – AIAS) như một mô hình thực tiễn, linh hoạt để tích hợp đạo đức trí tuệ nhân tạo tạo sinh (GenAI) vào đánh giá giáo dục đại học. Các tác giả lập luận rằng sự phát triển nhanh chóng của GenAI, đặc biệt từ ChatGPT năm 2022, đã tạo ra sự thay đổi mô hình trong giáo dục, mang lại cơ hội đổi mới sư phạm nhưng cũng đặt ra thách thức về tính toàn vẹn học thuật và đạo đức sử dụng. Thay vì tiếp cận cảm đoán hoặc coi GenAI chủ yếu như công cụ gian lận, AIAS khuyến khích tích hợp có trách nhiệm, tập trung vào việc hỗ trợ dạy học và phát triển kỹ năng phê phán. Khung này bao gồm năm cấp độ tích

lũy, cho phép giảng viên chọn mức độ sử dụng GenAI phù hợp với mục tiêu học tập, đảm bảo minh bạch, công bằng và thích ứng với chính sách của các cơ sở giáo dục. Năm cấp độ của AIAS được thiết kế theo cách tăng dần mức độ tương tác với GenAI, từ hạn chế đến hợp tác toàn diện, với trọng tâm là sự đánh giá phê phán của con người và minh bạch. Các cấp độ được trình bày chi tiết như sau:

Cấp độ 1 - Không AI: Nhiệm vụ được hoàn thành hoàn toàn không có sự hỗ trợ của AI, nhằm kiểm tra kiến thức, kỹ năng và hiểu biết cá nhân thuần túy. AI không được sử dụng ở bất kỳ giai đoạn nào. Ví dụ như thảo luận trực tiếp, ý tưởng nhóm trong lớp, hoặc vấn đáp dưới giám sát của giảng viên.

Cấp độ 2 - AI hỗ trợ tạo ý tưởng và cấu trúc: AI được phép dùng để tạo cấu trúc ý tưởng hay gợi ý cải thiện, nhưng sản phẩm cuối cùng không chứa nội dung AI trực tiếp. Ví dụ như sử dụng AI để gợi ý chủ đề nghiên cứu, dàn ý bài viết, hoặc ý tưởng nhóm. Cấp độ này khuyến khích khám phá ý tưởng rộng hơn nhưng nhấn mạnh sản phẩm cuối phải do con người tự tạo.

Cấp độ 3 - AI hỗ trợ chỉnh sửa: AI dùng để cải thiện độ rõ ràng, chất lượng ngôn ngữ hoặc cấu trúc của nội dung gốc do học viên tạo, nhưng không tạo nội dung mới. Phải nộp phiên bản gốc trong phụ lục để minh chứng. Ví dụ như chỉnh sửa ngữ pháp, từ vựng, hoặc cấu trúc câu; chỉnh sửa hình ảnh mà không tạo mới. Cấp độ này hữu ích cho học viên không phải tiếng mẹ đẻ.

Cấp độ 4 - AI hoàn thành nhiệm vụ, đánh giá con người: AI được yêu cầu dùng để hoàn thành phần cụ thể của nhiệm vụ, nhưng học viên phải đánh giá phê phán đầu ra AI về tính liên quan, chính xác, thiên kiến và phù hợp. Phải trích dẫn nội dung AI. Ví dụ như tạo nội dung AI rồi so sánh với nội dung con người, phê phán định kiến, hoặc tích hợp vào dự án lớn. Cấp độ này nhấn mạnh đến tư duy phản biện.

Cấp độ 5 - AI toàn diện: AI được dùng như “đồng hành” xuyên suốt nhiệm vụ, cho phép sáng tạo hợp tác mà không cần chỉ rõ nội dung AI. Tập trung vào mục tiêu học tập tổng thể và đạo đức sử dụng. Ví dụ như đồng sáng tạo nội dung với chủ đề rộng, khám phá đa công cụ GenAI, tạo sản phẩm hoàn chỉnh như phần mềm hoặc nghệ thuật. Cấp độ này chuẩn bị học viên cho môi trường chuyên nghiệp nơi GenAI là tiêu chuẩn.

Khung đánh giá trí tuệ nhân tạo (AIAS) nhấn mạnh việc chuyển từ cách tiếp cận cảm đoán sang tích hợp có trách nhiệm, nhằm duy trì tính toàn vẹn học thuật đồng thời thúc đẩy kỹ năng phê phán và sáng tạo [10]. Cụ thể, khung này phê phán việc coi nội dung do GenAI tạo ra như hình thức gian lận mặc định, thay vào đó ủng hộ tích hợp đạo đức thông qua minh bạch (trích dẫn nội dung AI) và trách nhiệm cá nhân, phù hợp với lập luận của bài viết này về chuyển từ “chống gian lận” sang “thiết kế minh chứng tư duy”.

Bốn mô hình đánh giá sau đây có thể được ánh xạ lên các cấp độ cao của AIAS, đặc biệt mô hình đánh giá

quá trình và hồ sơ minh chứng trong bài viết tương đồng với cấp độ 4 (Full AI hoặc AI Collaboration trong phiên bản chỉnh sửa [13]), nơi người học sử dụng GenAI như một cộng tác viên toàn diện để đồng sáng tạo đầu ra, nhưng vẫn duy trì trọng tâm vào kết quả học tập tổng thể và sử dụng đạo đức, mà không cần phân định rõ ràng nội dung do AI tạo ra [13, 14]. Ở cấp độ này, AI hỗ trợ hoàn thành nhiệm vụ cụ thể, nhưng người học bắt buộc phải đánh giá phê phán đầu ra về tính chính xác, thiên kiến và mức độ phù hợp, giống như yêu cầu sự tham gia phê phán sâu sắc với nội dung do AI tạo mà khung nhấn mạnh. Tương tự, nguyên tắc “minh bạch hóa hỗ trợ AI” và “quy trách nhiệm” trong bài viết có thể bổ sung cho yêu cầu minh bạch của AIAS, đặc biệt khi khung khẳng định rằng AIAS nên được sử dụng như một công cụ hướng dẫn thiết kế lại đánh giá thay vì hạn chế sử dụng GenAI thông qua phát hiện hoặc cấm đoán [10, 14]. Điều này giúp mở rộng ứng dụng khung AIAS vào các ngành khoa học xã hội và nhân văn bằng cách nhấn mạnh các dấu vết học tập cụ thể như nhật ký quyết định và vấn đáp, những yếu tố làm phong phú thêm khung đánh giá trong bối cảnh Việt Nam, nơi ưu tiên lập luận phê phán và minh chứng tư duy. Sự kết nối này hoàn toàn phù hợp với các khuyến nghị hướng vào quá trình học tập từ O’Sullivan và các cộng sự [5] như cần triển khai các phương pháp đánh giá hướng vào quá trình để đo lường chính bản thân quá trình học tập, bao gồm bản nháp, phần phản ánh và những cải tiến lặp lại, thay vì chỉ tập trung vào sản phẩm cuối cùng. Những nguyên tắc này khẳng định mạnh mẽ rằng việc tích hợp minh chứng quá trình cùng với giải trình không chỉ giúp bảo vệ tính hợp lệ của đánh giá, mà còn thúc đẩy phát triển các năng lực cốt lõi khó bị GenAI thay thế hoàn toàn trong tương lai.

Mô hình 1: Đánh giá dựa trên quá trình

Mô hình này đặc biệt hiệu quả trong các nhiệm vụ sau như bài luận học thuật, bài phản biện, bài tổng quan tài liệu, hoặc bài phân tích tình huống hay các học phần nhằm phát triển kỹ năng tư duy lập luận, viết học thuật, và phương pháp nghiên cứu. Ở đó, năng lực học tập được xem như quá trình phát triển liên tục, với điểm số phân bổ cho các mốc: xác định vấn đề, đề cương, bản nháp, phản hồi, sửa chữa, và tự đánh giá. GenAI được phép ở các bước hỗ trợ (ví dụ: gợi ý cấu trúc, chỉnh sửa ngôn ngữ, tinh chỉnh ý tưởng), nhưng bắt buộc cung cấp dấu vết tiến triển (tiên triển chỉnh sửa, nhật ký quyết định) và giải trình rõ ràng (bình luận phê phán về thay đổi, lý do chấp nhận/bác bỏ gợi ý AI). Quy trình triển khai:

- Giai đoạn khởi đầu: Cung cấp đề bài kèm tiêu chí đánh giá rõ ràng và quy tắc minh bạch sử dụng GenAI (các bước được phép, cách khai báo sử dụng AI theo mẫu chuẩn, ví dụ: nêu công cụ, prompt, mức độ chỉnh sửa).
- Nộp đề cương/luận điểm chính: Kèm 3-5 nguồn tài liệu dự kiến, lý do lựa chọn từng nguồn (đánh

giá độ tin cậy, liên quan), và bảng so sánh nguồn nếu cần.

- Nộp bản nháp đầu tiên: Kèm “bảng tự kiểm” (khẳng định - bằng chứng - lập luận), tự nhận diện điểm còn yếu, và khai báo phần hỗ trợ từ AI (nếu có).
- Giai đoạn phản hồi và sửa chữa: Giảng viên hoặc bạn học cung cấp phản hồi; người học lập kế hoạch sửa đổi cụ thể (chỉ rõ thay đổi, lý do, bài học rút ra từ phản hồi và AI nếu dùng).
- Nộp sản phẩm hoàn thiện: Kèm bản giải trình sửa đổi (chỉ tiết các thay đổi, lý do chấp nhận/chỉnh sửa/bác bỏ gợi ý AI, bài học rút ra), phiên bản gốc (nếu AI hỗ trợ editing), và khai báo đầy đủ sử dụng GenAI.

Thang đánh giá gợi ý (sử dụng 4 tiêu chí chính, mỗi tiêu chí thang 4 mức: 4 = Xuất sắc, 3 = Đạt, 2 = Cần cải thiện, 1 = Chưa đạt; với trọng số khác nhau, ví dụ minh bạch học thuật chiếm tỷ lệ cao hơn để nhấn mạnh trách nhiệm):

- Diễn giải nhiệm vụ và đặt vấn đề (20%): Rõ ràng trọng tâm, đúng phạm vi yêu cầu, thể hiện sự hiểu sâu vấn đề.
- Chất lượng lập luận và bằng chứng (30%): Logic chặt chẽ, có phản biện hiệu quả, liên kết nguồn tài liệu hợp lý và đa dạng.
- Năng lực chỉnh sửa dựa trên phản hồi (25%): Có lý do rõ ràng cho các thay đổi, thể hiện sự cải thiện thực chất và học hỏi từ quá trình.
- Minh bạch học thuật (25%): Trích dẫn đúng chuẩn, khai báo sử dụng GenAI đầy đủ và chính xác, kiểm chứng thông tin đáng tin cậy, sẵn sàng giải trình lựa chọn.

Mô hình 2: Đánh giá theo sản phẩm có minh chứng

Mô hình này hiệu quả nhất trong các ngữ cảnh sau như môn học kéo dài 12-15 tuần, cần đánh giá sự tiến bộ liên tục thay vì chỉ kết quả cuối kỳ hay học phần có nhiều nhiệm vụ ứng dụng thực tiễn, chẳng hạn như đọc và phản hồi tài liệu, phân tích dữ liệu, dự án nhỏ, hoặc các hoạt động tích lũy kỹ năng (ví dụ: viết học thuật, nghiên cứu xã hội, phân tích dữ liệu). Khi đó, hồ sơ học tập (portfolio) coi học tập như sự tích lũy minh chứng đa dạng, với trọng tâm là phản tư sâu sắc, tiên bộ cá nhân và trách nhiệm học thuật thay vì chỉ sản phẩm đơn lẻ. Quy trình triển khai:

- Thiết kế khung hồ sơ học tập: Xây dựng cấu trúc với 6-8 mảnh minh chứng bắt buộc (ví dụ: bản phản hồi đọc, phân tích dữ liệu, bài viết ngắn, bản đồ lập luận, nhật ký quyết định) và 2-3 mảnh tự chọn để khuyến khích sáng tạo. Mỗi mảnh cần có mô tả mục tiêu rõ ràng, tiêu chí đánh giá cụ thể, và yêu cầu khai báo sử dụng GenAI (nếu có: công cụ, prompt chính, mức độ chỉnh sửa, lý do sử dụng).

- Thu thập và xây dựng Hồ sơ học tập: Người học nộp dần các mảnh minh chứng trong suốt học kỳ, kèm ghi chú ngắn về quá trình thực hiện, vai trò của GenAI (nếu dùng) và tự đánh giá ban đầu.
- Tổng thuật phản tư: Cuối kỳ, người học nộp bài tổng hợp (khoảng 1000-1500 từ hoặc tương đương đa phương tiện) trả lời các câu hỏi hướng dẫn: “Tôi đã học được gì từ quá trình? Tiên bộ ở đâu so với đầu kỳ? Minh chứng nào chứng minh điều đó? GenAI đã hỗ trợ như thế nào, tôi đã kiểm soát và phê phán nó ra sao? Những giới hạn/thiên kiến nào tôi nhận ra và khắc phục?”.
- Vấn đáp kiểm chứng ngắn (5-7 phút): Người học trình bày và giải thích 2-3 lựa chọn quan trọng nhất trong Hồ sơ (ví dụ: tại sao chọn mảnh này, cách GenAI ảnh hưởng đến quá trình, bài học rút ra từ phản tư), giúp giảng viên xác nhận tính xác thực, chiều sâu tư duy và trách nhiệm cá nhân.

Thang đánh giá gợi ý (sử dụng 4 tiêu chí chính, mỗi tiêu chí thang 4 mức: 4 = Xuất sắc, 3 = Đạt yêu cầu, 2 = Cần cải thiện, 1 = Chưa đạt; với trọng số khác nhau, ví dụ năng lực phản tư chiếm 30% để nhấn mạnh minh chứng tư duy):

- Độ phù hợp của minh chứng với mục tiêu học phần (20%): Các mảnh minh chứng liên quan trực tiếp, không chỉ sưu tầm cho đủ số lượng mà thể hiện sự lựa chọn có chủ đích và đa dạng.
- Chất lượng tư duy thể hiện trong minh chứng (25%): Lập luận logic chặt chẽ, đối chiếu nguồn tài liệu, kiểm chứng thông tin đáng tin cậy, và phản biện sâu sắc.
- Mức độ tiến bộ (25%): Có bằng chứng rõ ràng so sánh đầu kỳ và cuối kỳ (qua các mảnh minh chứng), thể hiện sự phát triển kỹ năng thực chất (ví dụ: từ lập luận đơn giản đến phức tạp, tích hợp phê phán AI hiệu quả).
- Năng lực phản tư và trách nhiệm học thuật (30%): Phản ánh tự đánh giá chân thực, minh bạch sử dụng GenAI (khai báo rõ ràng, trích dẫn đúng chuẩn), kiểm chứng thông tin, và sẵn sàng bảo vệ lựa chọn trong vấn đáp.

Mô hình 3: Đánh giá năng lực lập luận

Mô hình này đặc biệt hiệu quả trong các ngữ cảnh sau như các bài luận cuối kỳ hoặc nhiệm vụ có trọng số cao, nơi cần xác nhận tính xác thực và chiều sâu tư duy hay bài nghiên cứu nhỏ, đề xuất nghiên cứu, phân tích chính sách, phân tích văn bản, hoặc các nhiệm vụ đòi hỏi lập luận phê phán cao (ví dụ: trong khoa học xã hội, nhân văn, hoặc chính sách công). Mô hình này phù hợp với tinh thần AIAS ở cấp độ 1 (Không AI) và cấp độ 4 (AI toàn diện), nơi GenAI có thể hỗ trợ chuẩn bị (như gợi ý ý tưởng ở cấp độ thấp hơn) nhưng không thể thay thế được khả năng giải trình trực tiếp và làm chủ lập luận qua đối thoại sống động. Bảo vệ vấn đáp trực tiếp buộc người học

phải chứng minh sự làm chủ tư duy, giảm thiểu rủi ro “AI viết thay” bằng cách tập trung vào khả năng diễn đạt tự nhiên, phản biện tại chỗ và bảo vệ lựa chọn cá nhân. Quy trình triển khai:

- Nộp sản phẩm ban đầu: Người học nộp bài viết, đề cương hoặc báo cáo kèm theo 1 trang tóm tắt luận điểm và bằng chứng chính để làm cơ sở cho phiên vấn đáp. Khai báo minh bạch sử dụng GenAI (nếu có: công cụ, mục đích hỗ trợ, mức độ chỉnh sửa) để giảng viên chuẩn bị câu hỏi phù hợp.
- Chuẩn bị câu hỏi: Giảng viên chuẩn bị 3 - 5 câu hỏi theo ba tầng độ phức tạp: Tầng 1 - Hiểu: Xác nhận nắm bắt nội dung cơ bản và khái niệm chính; Tầng 2 - Phân tích: Khai thác mối quan hệ giữa ý tưởng, bằng chứng, logic lập luận và nguồn tài liệu; Tầng 3 - Phản biện: Đặt câu hỏi thách thức, yêu cầu bảo vệ lập luận trước phản biện, mở rộng ý tưởng hoặc thừa nhận giới hạn.
- Thực hiện vấn đáp: Phiên vấn đáp kéo dài 8 -12 phút (có thể theo nhóm nhỏ 2 - 3 người để tiết kiệm thời gian và tăng tính tương tác), với các yêu cầu chính: giải thích luận điểm bằng ngôn ngữ tự nhiên của bản thân (không đọc thuộc lòng hoặc dựa vào văn bản AI); nêu rõ điểm yếu trong lập luận, cách củng cố và bài học rút ra từ quá trình; phản hồi một phản biện giả định hoặc câu hỏi mới từ giảng viên, thể hiện tư duy linh hoạt và phê phán.
- Đánh giá: Chấm điểm dựa hoàn toàn vào chất lượng thể hiện trong phiên vấn đáp (không dựa vào việc “đoán” nguồn gốc AI của sản phẩm viết). Có thể ghi âm/ghi hình (với sự đồng ý của người học) để lưu trữ, đảm bảo công bằng và minh bạch.

Thang đánh giá gợi ý (sử dụng 4 tiêu chí chính, mỗi tiêu chí thang 4 mức: 4 = Xuất sắc, 3 = Đạt yêu cầu, 2 = Cần cải thiện, 1 = Chưa đạt; trọng số có thể điều chỉnh, ví dụ trách nhiệm học thuật chiếm 30% để nhấn mạnh minh bạch và làm chủ):

- Hiểu nội dung và diễn giải rõ ràng (20%): Thể hiện sự nắm bắt sâu, diễn đạt tự nhiên bằng lời nói của mình (không học thuộc hoặc đọc thuộc lòng từ văn bản), giải thích khái niệm mạch lạc.
- Lập luận và phản biện (30%): Trả lời thuyết phục các câu hỏi “vì sao”, “dựa vào đâu”, với lập luận logic chặt chẽ, bằng chứng phù hợp, khả năng phản biện hiệu quả và liên kết ý tưởng.
- Khả năng xử lý câu hỏi mới (25%): Tư duy linh hoạt tại chỗ, không né tránh, xử lý tốt các tình huống bất ngờ hoặc thách thức, mở rộng lập luận khi cần.
- Trách nhiệm học thuật (25%): Thừa nhận giới hạn của lập luận, minh bạch về hỗ trợ từ GenAI (nếu có, giải trình cách kiểm soát và phê phán AI), sẵn sàng bảo vệ các lựa chọn cá nhân một cách đạo

đức và có cơ sở.

kiểm chứng tính xác thực, chiều sâu tư duy và trách nhiệm cá nhân.

Mô hình 4: Đánh giá dự án có nhật ký quyết định

Mô hình này đặc biệt hiệu quả trong các ngữ cảnh sau như các môn học phương pháp nghiên cứu, phân tích chính sách, truyền thông, văn hóa, xã hội học ứng dụng, hoặc các lĩnh vực đòi hỏi gắn kết lý thuyết với thực tiễn hay học phần cần người học làm việc với dữ liệu thực địa, tình huống cụ thể, hoặc dự án ứng dụng (ví dụ: phân tích chính sách công, nghiên cứu cộng đồng, thiết kế truyền thông xã hội). Thành phần nhật ký quyết định (decision log) đảm bảo sự đánh giá của con người (human evaluation), tính truy vết và trách nhiệm cá nhân, từ đó giảm thiểu rủi ro lạm dụng thụ động GenAI bằng cách ghi lại các quyết định then chốt, lý do, bằng chứng và bài học rút ra. Quy trình triển khai:

- Chọn vấn đề và đặt câu hỏi nghiên cứu: Người học nộp 1 trang đề xuất vấn đề, câu hỏi dự án, và lý do chọn đề tài (có thể kèm khai báo nếu dùng GenAI để gợi ý ý tưởng ban đầu).
- Lập kế hoạch dự án: Xây dựng timeline, danh sách nguồn dữ liệu dự kiến, nhận diện rủi ro tiềm ẩn, phương pháp kiểm chứng thông tin, và kế hoạch sử dụng GenAI (nếu áp dụng) với lý do minh bạch.
- Thu thập và khai thác dữ liệu: Thực hiện thu thập dữ liệu thực địa hoặc thứ cấp, tuân thủ quy tắc ghi nguồn, đạo đức nghiên cứu và kiểm chứng tính xác thực (ví dụ: đối chiếu nguồn, nhận diện thiên kiến nếu dùng AI hỗ trợ phân tích).
- Nộp bản nháp phân tích: Kèm theo ít nhất 5 quyết định quan trọng nhất đã thực hiện đến thời điểm này (ví dụ: thay đổi khung lý thuyết vì lý do dữ liệu không phù hợp, điều chỉnh phương pháp vì hạn chế nguồn, hoặc bác bỏ gợi ý AI vì thiên kiến), kèm nhật ký quyết định sơ bộ.
- Hoàn thiện sản phẩm cuối cùng: Tạo báo cáo, áp phích học thuật, báo cáo tóm tắt chính sách, video trình bày, hoặc sản phẩm khác phù hợp với mục tiêu dự án.
- Nộp nhật ký quyết định đầy đủ cộng với khai báo AI và trình bày ngắn:
 - Nhật ký quyết định: Tổng hợp tất cả các quyết định then chốt (ít nhất 8 - 10 mục), kèm lý do thuyết phục, bằng chứng hỗ trợ (dữ liệu, nguồn, phản hồi), vai trò của GenAI (nếu có), và bài học từ sai sót hoặc điều chỉnh.
 - Bản khai báo sử dụng GenAI: Nêu rõ bước nào dùng AI, mục đích cụ thể, câu lệnh chính (nếu phù hợp), mức độ chỉnh sửa sau AI, và cách phê phán/kiểm soát đầu ra.
 - Trình bày ngắn (7 phút): Người học trình bày dự án, nhấn mạnh các quyết định quan trọng và nhật ký, trả lời 2-3 câu hỏi từ giảng viên để

Thang đánh giá gợi ý (sử dụng 4 tiêu chí chính, mỗi tiêu chí thang 4 mức: 4 = Xuất sắc, 3 = Đạt yêu cầu, 2 = Cần cải thiện, 1 = Chưa đạt; trọng số có thể điều chỉnh, ví dụ chất lượng nhật ký quyết định chiếm 30% để nhấn mạnh tính truy vết và đánh giá mang tính nhân văn):

- Chất lượng câu hỏi nghiên cứu và tính phù hợp của phương pháp (20%): Câu hỏi rõ ràng, có ý nghĩa thực tiễn; phương pháp được chọn phù hợp với mục tiêu, bối cảnh và khả năng thực hiện.
- Lập luận dựa trên dữ liệu/bằng chứng (25%): Phân tích logic chặt chẽ, sử dụng dữ liệu thực chất (không chỉ dựa vào ý kiến cá nhân hoặc AI thô), có đối chiếu nguồn, kiểm chứng thông tin và nhận diện thiên kiến.
- Chất lượng nhật ký quyết định (30%): Quyết định được ghi chép rõ ràng, có lý do thuyết phục, bằng chứng hỗ trợ, thể hiện quá trình học hỏi từ sai sót, điều chỉnh hợp lý và vai trò phê phán GenAI (nếu dùng).
- Trách nhiệm học thuật (25%): Trích dẫn nguồn đầy đủ và đúng chuẩn, minh bạch sử dụng GenAI (khai báo chi tiết, giải trình cách kiểm soát), kiểm chứng thông tin đáng tin cậy, và sẵn sàng bảo vệ các lựa chọn trong phần trình bày.

Bốn mô hình đánh giá nêu trên không phủ nhận giá trị của văn bản như một yếu tố quan trọng trong quá trình học tập, mà ngược lại, đặt văn bản vào một hệ thống minh chứng toàn diện và đa chiều hơn. Trong hệ thống này, văn bản không còn là mục tiêu cuối cùng hay bằng chứng duy nhất, mà chỉ là một phần trong chuỗi tiến trình, lựa chọn và giải trình, nơi người học buộc phải thể hiện rõ ràng năng lực tư duy thực chất thông qua các dấu vết cụ thể như bản nháp, nhật ký quyết định, phần viết phản tư, vấn đáp, và khai báo sử dụng AI. Cách tiếp cận này giúp loại bỏ tình trạng phải phán đoán về nguồn gốc sản phẩm, đồng thời giảm đáng kể rủi ro lạm dụng GenAI một cách thụ động, không có trách nhiệm.

Trong bối cảnh GenAI phát triển nhanh chóng và mạnh mẽ, hướng đi này thực tế và bền vững hơn nhiều so với việc tập trung toàn bộ nguồn lực vào việc truy vết “ai là người viết văn bản này”, một nỗ lực ngày càng tốn kém, khó khăn và dễ rơi vào vòng luẩn quẩn giữa kiểm soát chặt chẽ và các hành vi đối phó tinh vi. Thay vì vậy, việc chuyển trọng tâm sang minh chứng tư duy cùng với trách nhiệm cá nhân của người học sẽ phù hợp hơn với sứ mệnh cốt lõi của giáo dục đại học: rèn luyện khả năng lập luận chặt chẽ, phán đoán có cơ sở vững chắc, và chịu trách nhiệm học thuật một cách đầy đủ, minh bạch.

Kết luận

Sự phổ biến của trí tuệ nhân tạo tạo sinh đã đặt ra một thách thức mang tính cấu trúc đối với hệ thống kiểm tra - đánh giá ở giáo dục đại học: đầu ra văn bản không còn là chỉ báo đáng tin cậy để suy ra năng lực tư duy học thuật. Trong bối cảnh đó, các biện pháp dựa trên cảm đoán hoặc truy vết nguồn gốc văn bản có thể tạo cảm giác kiểm soát tạm thời, nhưng khó giải quyết tận gốc vấn đề về tính hợp lệ của đánh giá, đồng thời dễ dẫn đến các hành vi đối phó và làm lệch lạc trọng tâm giáo dục khỏi việc phát triển năng lực thực chất. Vì vậy, hướng đi bền vững không nằm ở việc loại bỏ trí tuệ nhân tạo khỏi lớp học, mà ở việc tái thiết kế đánh giá để đo lường chính xác những năng lực cốt lõi mà giáo dục đại học cần bồi dưỡng. Để duy trì tính hợp lệ của việc đánh giá, cần một chuyển hệ mục tiêu từ đánh giá chủ yếu dựa vào đầu ra văn bản sang đánh giá dựa vào quá trình tư duy và minh chứng học tập. Trong hệ mục tiêu mới, văn bản không bị loại bỏ, mà được đặt trong một hệ bằng chứng rộng lớn hơn, nơi người học phải bộc lộ năng lực thực chất thông qua các dấu vết có thể truy vết và giải trình được, chẳng hạn như lựa chọn học thuật, đối chiếu nguồn, phản tư và giải trình. Cách tiếp cận này chuyển trọng tâm từ việc cố gắng xác định “ai viết” sang việc đo lường “người học đã tư duy và học như thế nào”. Qua đó, vấn đề cốt lõi không phải là “có dùng trí tuệ nhân tạo hay không”, mà là dùng ở mức độ nào, trong điều kiện nào, và với tiêu chí nào để vẫn đảm bảo toàn vẹn học thuật.

Để các mô hình không dừng lại ở lý thuyết mà vận hành hiệu quả ở giáo dục đại học Việt Nam, cần hội đủ bốn điều kiện then chốt. Thứ nhất, phải hoàn thiện khung chính sách - thể chế ở cả cấp Bộ và cơ sở giáo dục, công nhận giá trị tương đương giữa đánh giá quá trình, vấn đáp, hồ sơ minh chứng với bài thi cuối kỳ; đồng thời phân biệt rõ sử dụng AI có khai báo với gian lận học thuật. Thứ hai, nâng cao năng lực và động lực cho giảng viên thông qua việc bồi dưỡng ba kỹ năng cốt lõi bao gồm thiết kế đánh giá theo nguyên tắc căn chỉnh xây dựng (constructive alignment), năng lực phản hồi chất lượng cao (feedback literacy), và năng lực AI đủ để đặt câu hỏi vấn đáp phân biệt tư duy thật của sinh viên với nội dung do AI tạo ra. Thứ ba, giải quyết hạn chế về quy mô lớp học bằng cách thí điểm trước ở các học phần chuyên ngành có sĩ số dưới 40 sinh viên, kết hợp vấn đáp nhóm nhỏ, trợ giảng và đánh giá đồng cấp. Thứ tư, chuyển đổi văn hóa đánh giá từ nặng về điểm số tổng kết sang coi trọng quá trình học tập, đòi hỏi sự thay đổi nhận thức đồng bộ ở sinh viên, giảng viên và nhà quản lý, với khoảng thời gian thích hợp 2-3 học kỳ để mô hình phát huy hiệu quả.

Lời cảm ơn

Không

Phê duyệt đạo đức và sự đồng thuận tham gia

Không áp dụng

Nguồn tài trợ

Không

Xung đột lợi ích

Không có xung đột lợi ích nào được công bố

Tài liệu

- [1] C. Bryan and K. Clegg, *Innovative Assessment in Higher Education: A Handbook for Academic Practitioners*. London, UK: Routledge, 2019.
- [2] M. S. Ibarra-Sáiz, G. Rodríguez-Gómez, and D. Boud, “The quality of assessment tasks as a determinant of learning,” *Assess. Eval. High. Educ.*, vol. 46, no. 6, pp. 943–955, 2021, doi: [10.1080/02602938.2020.1828268](https://doi.org/10.1080/02602938.2020.1828268).
- [3] Q. Xia, X. Weng, F. Ouyang, T. J. Lin, and T. K. Chiu, “A scoping review on how generative artificial intelligence transforms assessment in higher education,” *Int. J. Educ. Technol. High. Educ.*, vol. 21, no. 1, p. 40, 2024, doi: [10.1186/s41239-024-00468-z](https://doi.org/10.1186/s41239-024-00468-z).
- [4] Z. N. Khlaif, W. A. Alkhouk, N. Salama, and B. A. Eideh, “Redesigning assessments for AI-enhanced learning: A framework for educators in the generative AI era,” *Educ. Sci.*, vol. 15, no. 2, p. 174, 2025, doi: [10.3390/educsci15020174](https://doi.org/10.3390/educsci15020174).
- [5] J. O’Sullivan, C. Lowry, R. Woods, and T. Conlon, “Generative AI in higher education teaching & learning: Principles for ethical AI adoption,” Higher Education Authority, Tech. Rep., 2025, doi: [10.82110/qmt6-jw48](https://doi.org/10.82110/qmt6-jw48).
- [6] T. K. F. Chiu, “The impact of generative AI (GenAI) on practices, policies and research direction in education: a case of ChatGPT and Midjourney,” *Interact. Learn. Environ.*, vol. 32, no. 10, pp. 6187–6203, 2024, doi: [10.1080/10494820.2023.2253861](https://doi.org/10.1080/10494820.2023.2253861).
- [7] E. Kasneci *et al.*, “ChatGPT for good? on opportunities and challenges of large language models for education,” *Learn. Individ. Differ.*, vol. 103, p. 102274, 2023, doi: [10.1016/j.lindif.2023.102274](https://doi.org/10.1016/j.lindif.2023.102274).
- [8] K. Rajaram, *Learning Intelligence: Innovative and Digital Transformative Learning Strategies – Cultural and Social Engineering Perspectives*. Singapore: Springer Nature Singapore, 2023.
- [9] G. Makransky, S. Borre-Gude, and R. E. Mayer, “Motivational and cognitive benefits of training in immersive virtual reality based on multiple

- assessments," *J. Comput. Assist. Learn.*, vol. 35, no. 6, pp. 691–707, 2019, doi: [10.1111/jcal.12375](https://doi.org/10.1111/jcal.12375).
- [10] M. Perkins, L. Furze, J. Roe, and J. MacVaugh, "The artificial intelligence assessment scale (AIAS): A framework for ethical integration of generative AI in educational assessment," *J. Univ. Teach. Learn. Pract.*, vol. 21, no. 6, 2024, doi: [10.53761/q3azde36](https://doi.org/10.53761/q3azde36).
- [11] D. Sallai, J. Cardoso-Silva, M. Barreto, F. Panero, G. Berrada, and S. Luxmoore, "Approach generative AI tools proactively or risk bypassing the learning process in higher education," *LSE Public Policy Rev.*, vol. 3, no. 3, p. 7, 2024, doi: [10.31389/lseppr.108](https://doi.org/10.31389/lseppr.108).
- [12] E. Page, G. Meyers, and E. K. Billings, "Theory to practice: An assessment framework for generative AI," *Intersection: J. Intersect. Assess. Learn.*, vol. 5, no. 4, pp. 114–126, 2024, doi: [10.61669/001c.124250](https://doi.org/10.61669/001c.124250).
- [13] M. Perkins, J. Roe, and L. Furze, "The AI assessment scale revisited: A framework for educational assessment," *arXiv preprint arXiv:2412.09029*, 2024, doi: [10.48550/arXiv.2412.09029](https://doi.org/10.48550/arXiv.2412.09029).
- [14] —, "Reimagining the artificial intelligence assessment scale (AIAS): A refined framework for educational assessment," *J. Univ. Teach. Learn. Pract.*, vol. 22, no. 7, 2025, doi: [10.53761/rrm4y757](https://doi.org/10.53761/rrm4y757).
- [15] T. Corbin, J. Walton, P. Bannister, and J. P. Deranty, "On the essay in a time of GenAI," *Educ. Philos. Theory*, vol. 58, no. 3, pp. 198–210, 2026, doi: [10.1080/00131857.2025.2572802](https://doi.org/10.1080/00131857.2025.2572802).
- [16] B. Exintaris, N. Karunaratne, and E. Yuriev, "Metacognition and critical thinking: Using ChatGPT-generated responses as prompts for critique in a problem-solving workshop (SMARTCHEMPer)," *J. Chem. Educ.*, vol. 100, no. 8, pp. 2972–2980, 2023, doi: [10.1021/acs.jchemed.3c00481](https://doi.org/10.1021/acs.jchemed.3c00481).
- [17] H. Vartiainen and M. Tedre, "Using artificial intelligence in craft education: crafting with text-to-image generative models," *Digit. Creat.*, vol. 34, no. 1, pp. 1–21, 2023, doi: [10.1080/14626268.2023.2174557](https://doi.org/10.1080/14626268.2023.2174557).